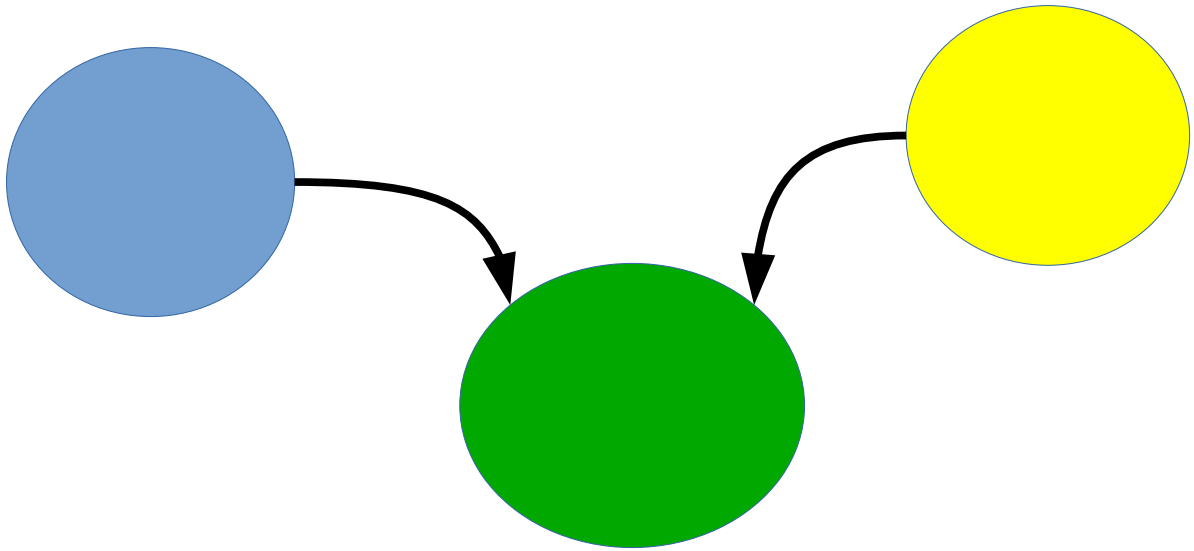
Bioinformatics
CS300
Chap 3
Sequence Alignment and an Influenza Outbreak

Week6, Deck 1
Fall 2022
Oliver BONHAM-CARTER

Descent with Modification

- ***Descent with modification*** is simply a passing trait from parent to offspring.
- One of the fundamental ideas behind Charles Darwin's theory of evolution.
- Traits are passed on to children in a process known as heredity.

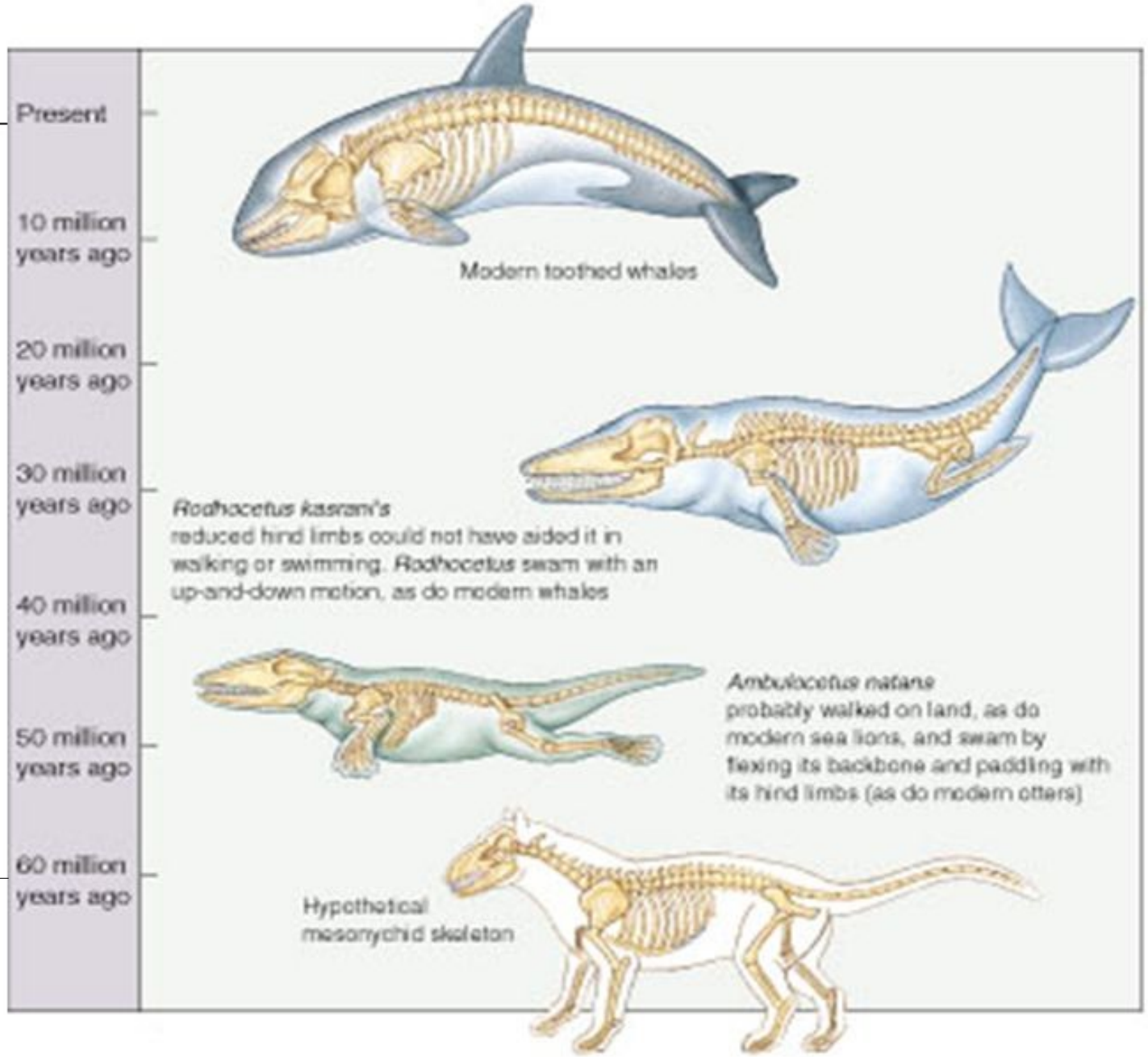


Passed Down From Genome to Genome

- DNA replication ensures a **mostly** faithful passing of the genome to progeny
- What would be the consequence of 100% accurate replication?
- Is that high similarity really desirable for a species?
- How does decent with modification happen?

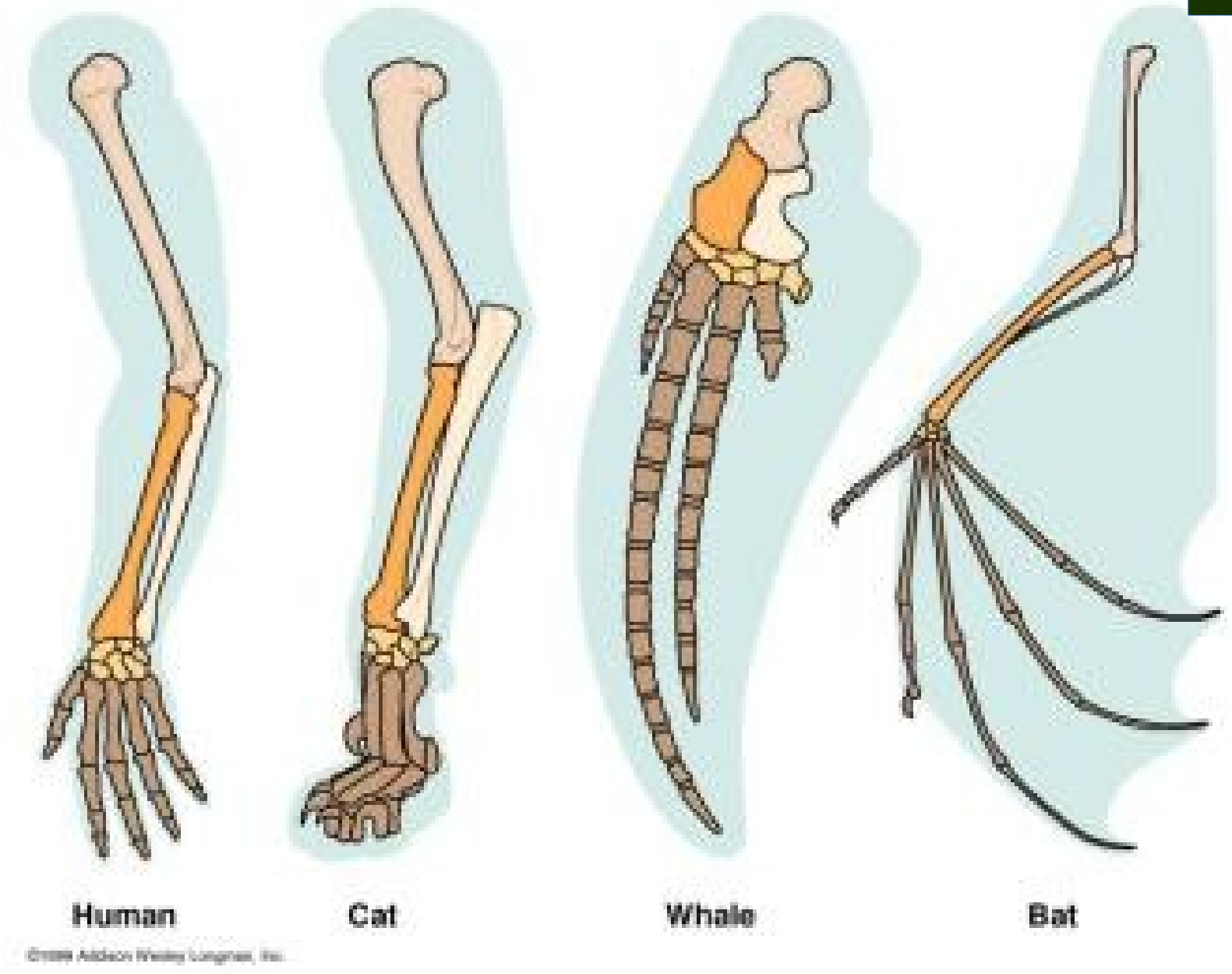
Descent With Modification

Present Day

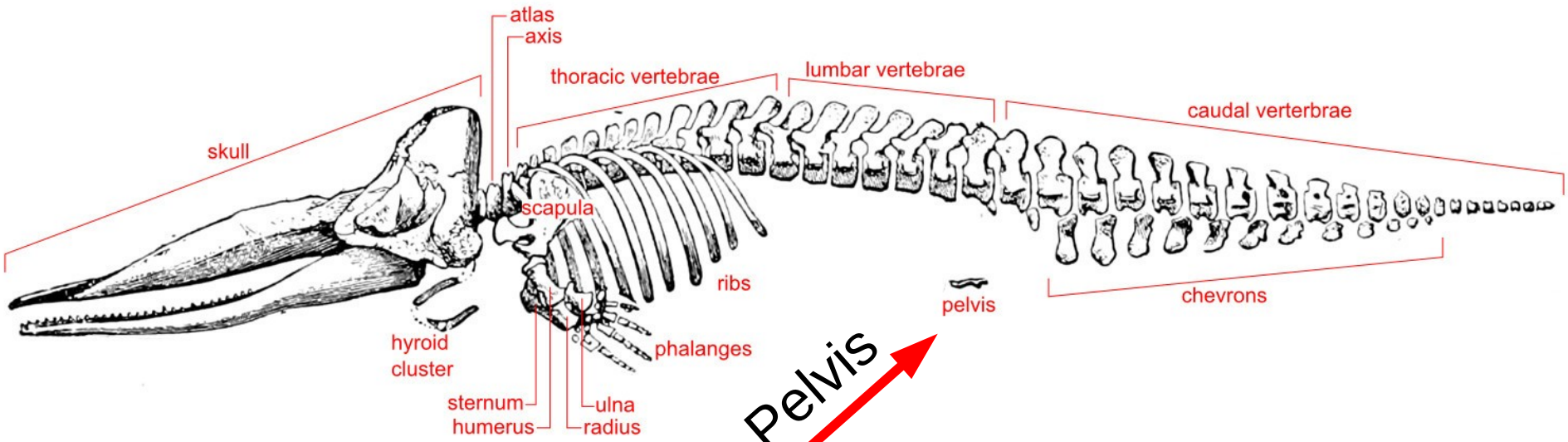


Long Ago

Same Bone, Different Day

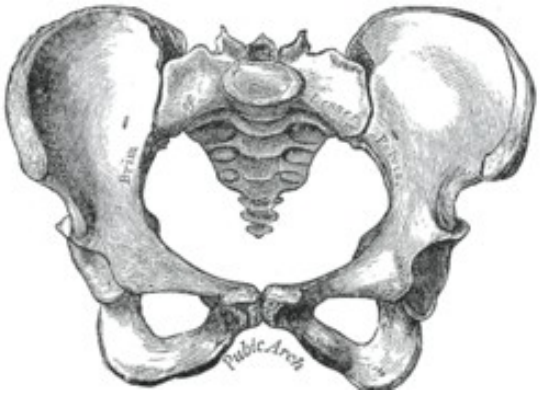


Same Bone, Different Day



Whale's Pelvis

Human's Pelvis



How Does Descent With Modification Happen?

Mutation

- A change in a DNA sequence
- Results from errors in replication or repair
- Mutation is the ultimate source of genetic variation

Domestic Dog – *Canis lupus familiaris*
All descendent from the grey wolf
All the same species (sub-species)
Breeds – variation within the species

Sequence Variations

- Sequences may have diverged from a common ancestor through various types of mutations:

- substitutions (A**C**GA → A**G**GA)
- insertions (ACGA → AC**CGGA**GA)
- deletions (A**CGGA**GA → AGA)

  - You are UNIQUE and SLIGHTLY GENETICALLY DIFFERENT from each of your parents, grand parents, great grand parents …
  - Retro Viruses (Influenza, HIV, etc) are also unique and slightly genetically different from their ancesters

Coding Time!

Research the code!

- Locate the supplied sequence alignment code in your `sandbox/` directory for the week.
- Make a virtual environment and play with the code. Instructions in the source code file.
- **Short comings**: The tool only processes two sequences at a time. *Can you fix this short coming, or add some improvement to the code?*
- Please add one of the sequence tests from your lab or from another source to the code to improve its ability to analyze sequences.
- When that has been completed, go on to the next part.

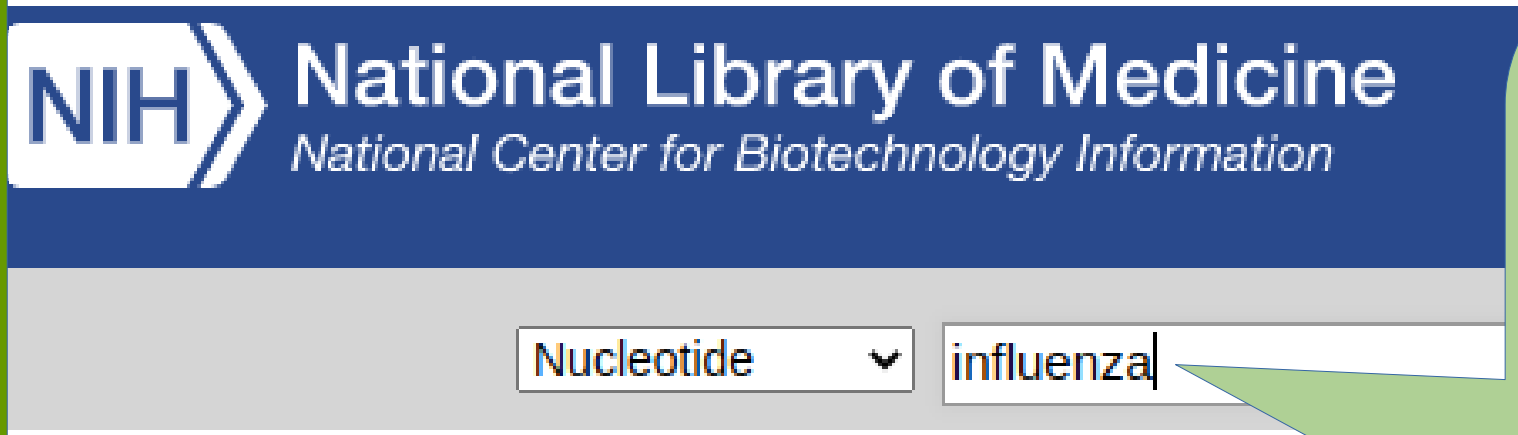
Research genetic diversity!

- Prove that **genetic drift** occurs in diseases such as *Influenza*.
- Go to https://www.ncbi.nlm.nih.gov/ to perform a search for sequences
- Choose ten closely related sequences from different organisms or origins.
- Run these sequences through the sequence analysis code to make a case that genetic drift happens.

You can work and present in groups. Each person is to submit a copy of the work for grading purposes.

Research genetic diversity!

ALLEGHENY COLLEGE



Try searching for *types of diseases; covid, sars, influenza, etc.* for example.

GitHub Classroom working repository:
https://classroom.github.com/a/SV1VA3A9

Friday:
Please come with two or three slides to present your improvements to the tool and to discuss your experiments.

Due on Friday

07 Oct 2022